Isolation, Composition and Ordering of Cyanogen Bromide Peptides in Human Casein

4282

By MERTON L. GROVES and RAE GREENBERG
**Eastern Regional Research Center*, 600 *East Mermaid Lane, Philadelphia, PA* 19118, *U.S.A.*

The major component of the casein fraction of human milk was cleaved by cyanogen bromide, and the composition of the resulting peptides was determined. Casein was also subjected to limited digestion by trypsin, and the amino acid composition of the isolated peptides was established. With this information the peptides were ordered as they occur in the purified protein.

The composition of the major component of the casein fraction in human milk resembles β-casein more closely than other cow's milk caseins. Unlike cow β-casein, which has a specific number of phosphate groups depending on the genetic type, human casein is comprised of multiphosphorylated forms with from zero to five phosphate groups per molecule (Groves & Gordon, 1970). The *N*-terminal sequence of the first 28 residues of this human casein has been determined, and the phosphorus in the fully phosphorylated form is found at residues 3, 6, 8, 9 and 10 (Greenberg *et al.*, 1976).

To sequence human β-casein further, its CNBr peptides and trypsin-cleaved lysine-blocked fragments have been isolated. Comparison of the composition of the two sets of peptides allows ordering the fragments in their proper alignment as they occur in the original molecule.

Experimental

Materials

Human caseins, I to VI, representing protein with from five to zero phosphate groups respectively, were isolated as described previously (Groves & Gordon, 1970).

Cleavage with CNBr and isolation of peptides

The CNBr cleavage of human casein was carried out essentially by the method of Phillips & Azari (1971). Approx. 100mg of casein was dissolved in 6ml of 70% (v/v) formic acid. To this was added 1.3ml of CNBr solution (1g in 70% formic acid), producing a 100-fold molar excess of reagent over methionine content. The reaction was allowed to proceed for 20h in an N_2 atmosphere at 25°C. The reaction mixture was then diluted tenfold with water and freeze-dried.

Initially the CNBr peptides were fractionated by gel filtration on Bio-Gel P30 (Bio-Rad Laboratories, Richmond, CA, U.S.A.) by using 20% (v/v) acetic acid on a 0.9cm×160cm column at a flow rate of 8ml/h. The *N*-terminal phosphopeptide fragment is insoluble in this solvent and was recovered in relatively pure form by centrifugation at low speed. To monitor the elution at 220nm, a solution of 0.1M-NH_4HCO_3 was also used with Bio-Gel P30 The phosphopeptide is readily soluble in this solvent. Peptides not separated by gel filtration were resolved by chromatography on Whatman DE-32 DEAE-cellulose with a 2cm×25cm column equilibrated with 5mM-NH_4HCO_3, pH8.3, 3°C. After regeneration of DEAE-cellulose in 0.5M-HCl followed by 0.5M-NaOH, it was suspended in 1M-NH_3, washed with water, suspended in 1M-NH_4HCO_3, adjusted to pH8.3, washed with 5mM-NH_4HCO_3 (adjusted to pH8.3), and then poured into a chromatographic column. The material was eluted at a flow rate of 12ml/h, with stepwise increase in NH_4HCO_3 concentration.

High-voltage electrophoresis of CNBr peptides

High-voltage electrophoresis on Whatman 3MM paper for 1.5h at 40V/cm in a horizontal water-cooled apparatus by using pyridine/acetic acid/water (25:1:225, by vol.), pH6.4, was used to purify two small CNBr peptides by the procedure of Ingram (1958). A sample (10mg) enriched in the small peptides was subjected to high-voltage electrophoresis; the peptides were identified by staining a narrow strip with ninhydrin [2,2-dihydroxy-1*H*-indene-1,3(2*H*)-dione]. With this as a guide, the peptides were then eluted from the remaining paper strips and recovered by freeze-drying.

Abbreviations used: peptides prefixed by 'CB-' are those produced by CNBr cleavage; those prefixed by 'T-' are produced by cleavage with trypsin.

* Agricultural Research Service, United States Department of Agriculture.

Trypsin cleavage of trifluoroacetylated casein and isolation of the peptides

S-Ethyl trifluoroacetate was used by the method of Goldberger (1967) to mask the lysine residues so that only the arginyl bonds were susceptible to trypsin. A suspension of 200mg of casein IV (two phosphate groups) in 15ml of water was placed in an autotitrator (Radiometer, Copenhagen, Denmark) and dissolved by titrating to pH9.0 with 2M-NaOH. Then 0.6ml of *S*-ethyl trifluoroacetate was added to the solution and the pH was maintained at 9.0 with 2.0M-NaOH. After 10min another 0.6ml of the reagent was added, and the solution was stirred for 1h. The trifluoroacetylated casein was precipitated at pH4.6 by the addition of dilute HCl. After centrifugation, the precipitate was washed with water, ethanol, acetone, and air-dried, with a resultant yield of 84%. In one experiment with casein V (one phosphate group), only 60% was recovered, indicating that the phosphorus content of the casein might affect the solubility of the derivative at acid pH.

For tryptic cleavage, the trifloroacetylated casein was placed in the autotitrator with water to give a 1% protein solution and was titrated with 0.1M-NaOH to pH8.3. A solution of trypsin treated with 1-chloro-4-phenyl-3-L-toluene-*p*-sulphonamidobutan-2-one ('TPCK') was added at a protein/enzyme ratio of 75:1 (w/w), and the digestion was carried out at pH8.3. The reaction proceeded at 25°C under N_2 for 1h, at which point the solution was heated for 5min in a boiling-water bath and the trifluoroacetylated fragments were recovered by freeze-drying.

To remove the blocking groups from the peptides, 2.5ml of 1M-piperidine at 3°C was added to 170mg of digest. The solution was placed in an ice/water bath for 2h and then acidified with 3.8ml of cold 1M-acetic acid. The solution was subjected to gel filtration on a column (2cm×71cm) of Sephadex G-25 (fine grade), at 3°C, equilibrated with 1M-acetic acid. Since the elution profile at 280nm indicated a lack of peak resolution, the eluate was divided into fractions designated A, B and C.

Phosphocellulose equilibrated with 0.01M-sodium acetate, pH4.0, was used to isolate the trypsin-cleaved peptides. Elution was effected by stepwise increase in molarity of this solution.

Amino acid composition

Procedures of Moore & Stein (1963) were used for automated amino acid analysis. Samples were hydrolysed at 110°C for 24h with 5.7M-HCl containing phenol (10μl/ml), in sealed evacuated tubes.

Polyacrylamide-gel disc electrophoresis

Gel-electrophoresis experiments were done at pH4.3 and 9.5 in 7.5% (w/v) polyacrylamide gels with or without urea as indicated in the Figures; staining was with Amido Black (Groves, 1974).

Results

CNBr peptides

Human casein contains three methionine residues and should produce four peptides on treatment with CNBr. Since the variable phosphorus region is in the *N*-terminal portion of the molecule and since methionine is absent from this area (Greenberg *et al.*, 1976), only the *N*-terminal CNBr cleaved peptide will show variations in mobility on electrophoresis that depend on the number of phosphate groups in the casein. Fig. 1 shows the separation of CNBr peptides of casein IV on Bio-Gel P30. Two large peptides, CB-1 and CB-4, are eluted first, followed by peptides CB-3, which contains no tyrosine, and finally peptide CB-2. The CB-2 and CB-3 peptides were each rechromatographed under the same conditions. Amino acid analyses of these two small column-purified peptides showed good agreement with corresponding peptides isolated by high-voltage electrophoresis.

The Bio-Gel P30 fraction containing CB-1 and CB-4 peptides was chromatographed on DEAE-cellulose (Fig. 2). The basic peptide, CB-4, was eluted with the starting buffer. After the change to 0.05M-buffer, the phosphopeptide CB-1 was eluted at NH_4HCO_3 concentration of about 0.1M.

Polyacrylamide-gel disc-electrophoretic patterns of the four CNBr peptides at alkaline and acid pH are

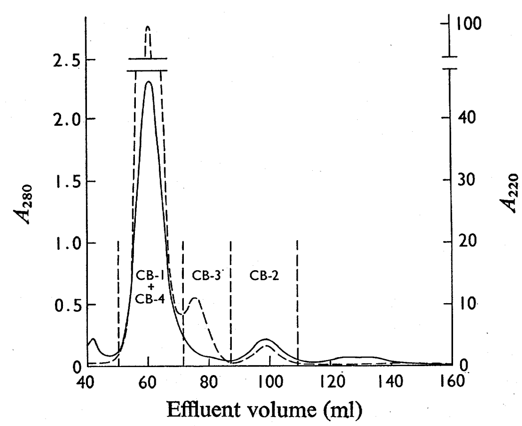

Fig. 1. *Gel filtration of CNBr peptides on Bio-Gel P*30
The column size was 0.9cm×160cm, and the flow rate 8ml/h; 60–100mg of sample was dissolved in 2–3ml of 0.1M-NH_4HCO_3. Vertical broken lines indicate the fractions that were combined; ——, A_{280}; ----, A_{220}.

phosphocellulose after removal of the blocking group and partial fractionation by gel filtration is shown in Fig. 4. This elution pattern shows the three fractions from the Sephadex G-25 column, which were chromatographed separately on phosphocellulose. The first Sephadex fraction (A) contains most of the largest peptide designated as peptide T-2. The phosphopeptide T-1A, insoluble in the starting buffer (0.01 M-sodium acetate, pH4.0), is isolated by centrifugation before chromatography of the peptides in the supernatant. Peptide T-1B′ differs on electrophoresis from peptide T-1B, but resembles T-1B in amino acid composition, except that it contains six fewer neutral amino acids. This is also true for peptides T-2′ and T-2, except that only minor differences in composition were observed. The T-1B′ and T-2′ peptides present in relatively small amounts may result from non-specific or incomplete tryptic cleavage. The mobility differences may result from amide hydrolysis due to the high pH of the reagent used to remove the trifluoroacetate group.

Polyacrylamide-gel disc-electrophoretic patterns of the trifluoroacetylated peptides are shown in Fig. 5. The gel for phosphopeptide T-1A at alkaline pH was overloaded, but shows two fast-moving zones at lower concentration. The reason for this heterogeneity is unknown; perhaps difference in the number of amide groups is a factor. The acidic phosphopep-

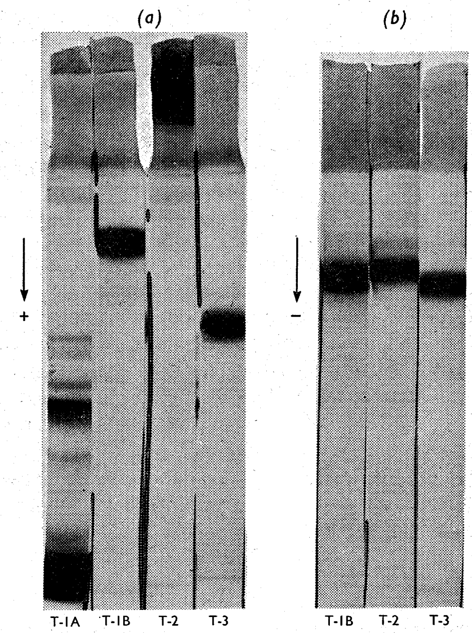

Fig. 5. *Disc-gel-electrophoretic patterns of tryptic peptides*
(*a*) Standard gels, pH9.5, 7.5% (w/v) acrylamide, 4M-urea; (*b*) pH4.3 gels.

Table 2. *Composition of peptides produced by trypsin cleavage*

The composition of human casein is taken from Groves & Gordon (1970); all other analyses represent 24 hydrolysates. Reported data are the averages of single analyses of two to four preparations of each peptide. Ratios determined as follows: T-1A, Lys≡4; T-1B, Lys≡3; T-2, Lys≡3; T-3, Asp≡3. Tryptophan was qualitatively determined by using Erlich's reagent; threonine and serine are uncorrected for losses. Numbers in parentheses are nearest-integer values.

Amino acid	Composition (residues/mol of peptide) Peptide T-1A	Peptide T-1B	Peptide T-2	Peptide T-3	Total T	Human β-casein
Lys	4.0 (4)	3.0 (3)	3.0 (3)	0.5 (1)	11	11
His	1.6 (2)	—	0.9 (1)	1.8 (2)	5	5
Arg	0.7 (1)	0.9 (1)	0.9 (1)	—	3	3
Asp	2.8 (3)	2.4 (2)	3.2 (3)	3.0 (3)	11	11
Thr	1.7 (2)	2.0 (2)	2.9 (3)	1.9 (2)	9	9
Ser	4.5 (5)	1.1 (1)	2.2 (2)	1.4 (1)	9	9
Glu	13.3 (13)	6.5 (7)	12.1 (12)	6.0 (6)	38	39
Pro	4.3 (4)	9.2 (9)	19.0 (19)	7.1 (7)	39	39
Gly	1.4 (1)	2.9 (3)	0.6 (1)	0.7 (1)	6	3
Ala	—	3.1 (3)	1.3 (1)	3.0 (3)	7	7
Val	2.3 (2)	5.3 (5)	6.5 (7)	3.7 (4)	18	19
Met	—	0.7 (1)	1.7 (2)	—	3	3
Ile	3.1 (3)	3.5 (3)	3.9 (4)	2.1 (2)	12	13
Leu	2.3 (2)	3.6 (4)	12.8 (13)	7.2 (7)	26	26
Tyr	2.2 (2)	2.5 (3)	1.1 (1)	1.0 (1)	7	7
Phe	1.1 (1)	1.7 (2)	1.9 (2)	—	5	5
Trp	—	—	1	—	1	1

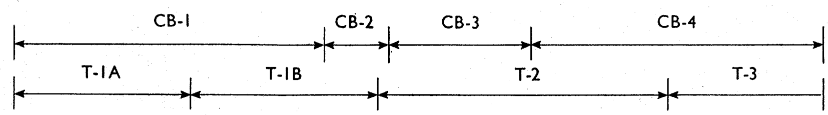

Fig. 6. *Order of the CNBr and tryptic peptides in human β-casein*
For details, see the text.

tide T-1A is not shown at pH4.3, since it does not move into the gel.

Human β-casein contains three arginine residues, one of which is *N*-terminal; thus three peptides would be expected after trypsin digestion of the trifluoroacetylated derivative. The amino acid composition of these peptides (Table 2) shows that four peptides are required to approximate the amino acids in the parent molecule. Possibly a lysine residue between peptides T-1A and T-1B is not protected or a non-specific cleavage by trypsin occurs.

Discussion

Examination of the relative composition of CNBr peptides in Table 1 reveals that peptide CB-1 is *N*-terminal, since it is the only peptide containing at least four serine residues, the number reported in the sequence of the first 28 residues (Greenberg *et al.*, 1976). Also peptide CB-4 is *C*-terminal, since it contains no homoserine lactone. The positions of peptides CB-2 and CB-3 can be deduced from information in Table 2. Peptide T-1A, like peptide CB-1, must be *N*-terminal because of the serine requirement mentioned above. Also, a summation of each amino acid in peptides T-1A plus T-1B shows that these two peptides account for all of peptide CB-1, with about 15 residues remaining. Comparison of these residues with the amino acids in peptides CB-2 and CB-3 shows that only peptide CB-2 fits, and that peptides T-1A plus T-1B contain most of the 16 residues in peptide CB-2. These conclusions can be confirmed by an examination of the arginine and alanine content of the peptides. The relatively limited numbers of these residues and their distribution dictates only one allowable configuration. By these and further comparisons of amino acid relationships between the two sets of peptides, the order (Fig. 6) in which the peptides occur in the original protein is deduced.

We thank Mr. H. J. Dower for performing the amino acid analyses. Reference to a brand or firm name in this paper does not constitute endorsement by the United States Department of Agriculture over others of a similar nature not mentioned.

References

Goldberger, R. F. (1967) *Methods Enzymol.* **11**, 317–322
Greenberg, R., Groves, M. L. & Peterson, R. F. (1976) *J. Dairy Sci.* **59**, 1016–1018
Groves, M. L. (1974) *Methods of Gel Electrophoresis of Milk Proteins* (Swaisgood, H., ed.), pp. 26–29, American Dairy Science Association (Nomenclature and Methodology of Milk Proteins Committee)
Groves, M. L. & Gordon, W. G. (1970) *Arch. Biochem. Biophys.* **140**, 47–51
Ingram, V. M. (1958) *Biochim. Biophys. Acta* **28**, 539–545
Moore, S. & Stein, W. H. (1963) *Methods Enzymol.* **7**, 819–831
Phillips, J. L. & Azari, P. (1971) *Biochemistry* **10**, 1160–1165

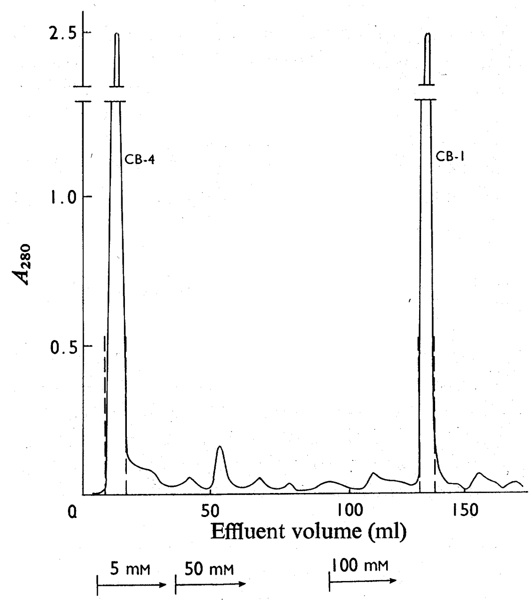

Fig. 2. *Chromatography of CNBr peptides CB-1 and CB-4 on DEAE-cellulose*

The column (2cm×25cm) was equilibrated with 5mM-NH_4HCO_3, pH8.3, 3°C. The column was eluted at a flow rate of 12ml/h, with stepwise increase in NH_4HCO_3 concentration. The sample, 70mg of peptide fraction suspended in 3ml of 5mM-NH_4HCO_3 adjusted to pH8.3, was dissolved at 3°C. Vertical broken lines indicate the fractions that were combined.

shown in Fig. 3. The phosphopeptide, CB-1, shows a double band at alkaline pH, perhaps owing to partial conversion of the *C*-terminal homoserine residue into homoserine lactone; at acid pH this peptide is not resolved. Peptide CB-4, which is shown below to be

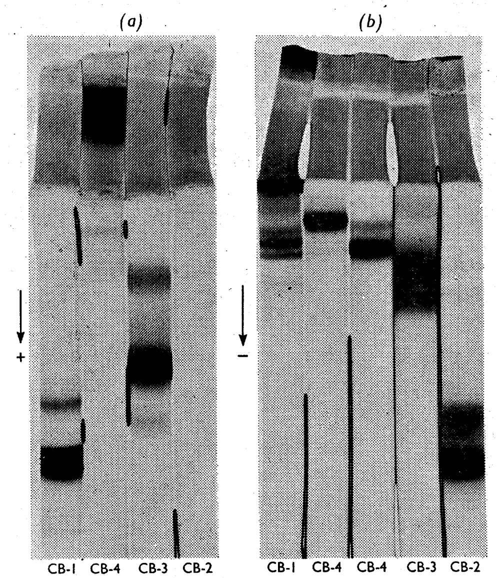

Fig. 3. *Disc-gel-electrophoretic pattern of CNBr peptides*
(*a*) standard gels, pH9.5, 7.5% (w/v) acrylamide; (*b*) pH4.3 gels, containing 8M-urea.

Table 1. *Composition of peptides produced by CNBr cleavage*

The composition of human casein is taken from Groves & Gordon (1970); all other analyses represent 24h hydrolysates. Reported data are the averages of single analyses of three to five preparations of each peptide. Ratios were determined by using Arg≡1, except for peptide CB-3, where His≡1. Methionine was determined qualitatively by the presence of homoserine lactone; tryptophan was determined qualitatively with Ehrlich's reagent; threonine and serine are uncorrected for losses. Numbers in parentheses are nearest-integer values.

	Composition (residues/mol of peptide)						
Amino acid	Peptide CB-1	Peptide CB-2	Peptide CB-3	Peptide CB-4	Sum of CB-peptides	CB-casein	Human β-casein
Lys	4.9 (5)	3.1 (3)	2.0 (2)	1.1 (1)	11	11	11
His	1.8 (2)	—	1.0 (1)	1.8 (2)	5	5	5
Arg	1.0 (1)	1.0 (1)	—	1.0 (1)	3	3	3
Asp	4.0 (4)	1.2 (1)	3.1 (3)	3.0 (3)	11	11	11
Thr	2.0 (2)	2.1 (2)	1.9 (2)	2.8 (3)	9	9	9
Ser	5.0 (5)	—	1.1 (1)	1.8 (2)	8	8	9
Glu	19.4 (19)	1.5 (1)	4.4 (4)	13.7 (14)	38	41	39
Pro	10.3 (10)	1.2 (1)	8.3 (8)	17.1 (17)	36	39	39
Gly	2.1 (2)	1.1 (1)	—	—	3	3	3
Ala	2.1 (2)	1.1 (1)	—	3.9 (4)	7	7	7
Val	5.3 (5)	3.3 (3)	1.6 (2)	8.9 (9)	19	19	19
Met	1	1	1	—	3	(3)	3
Ile	5.8 (6)	—	2.0 (2)	3.7 (4)	12	12	13
Leu	5.3 (5)	—	9.0 (9)	10.9 (11)	25	26	26
Tyr	3.8 (4)	1.0 (1)	—	1.9 (2)	7	7	7
Phe	2.6 (3)	—	1.7 (2)	—	5	5	5
Trp	—	—	—	1	1	1	1

the C-terminal peptide, did not enter the separating gel at alkaline pH. The two zones shown here at acid pH are the resolved slower and faster bands obtained by rechromatography of CB-4 peptide on DEAE-cellulose. The eluted peak was divided into three fractions, the first and last portion of the peak are shown (Fig. 3), whereas the middle fraction contains equal amounts of the slower and faster moving peptides. The amino acid composition of these two CB-4 peptides is the same. The mobility difference may result from differences in the amide content. Of the two smaller peptides, CB-2 and CB-3, only peptide CB-3 is found at alkaline pH; CB-2, a basic peptide does not move into the gel at this pH. At pH4.3, both peptides are present. When first eluted from the Bio-Gel, both peptides showed single electrophoretic zones; on fractionation and after storage at −20°C for several months the peptides show two components. This also might reflect the partial conversion of homoserine into homoserine lactone. Destaining the gels containing peptides CB-2 and CB-3 must be controlled to avoid diffusion of these small peptides from the gels.

The compositions of human β-casein, CNBr-treated casein and the CNBr peptides are shown in Table 1.

Peptides from trypsin-cleaved trifluoroacetylated casein

Human casein IV was treated with S-ethyl trifluoroacetate to produce trifluoroacetylated casein. On alkaline disc gel electrophoresis, trifluoroacetylated casein IV shows a single zone of faster mobility than that of the unmodified protein. Chromatography of the trypsin-cleaved trifluoroacetylated casein on

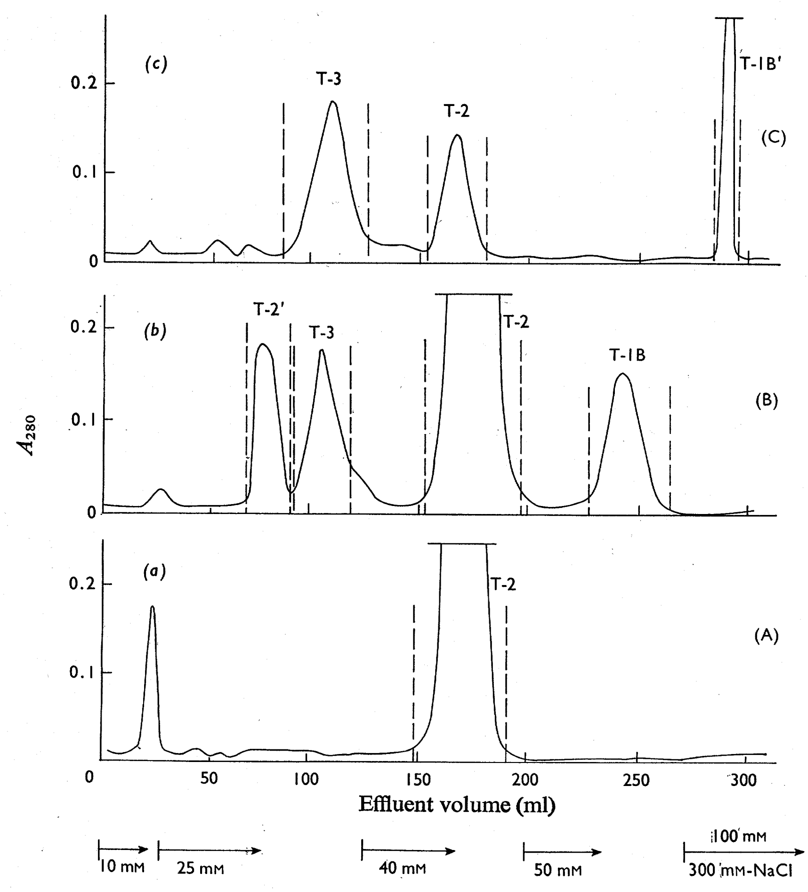

Fig. 4. *Phosphocellulose chromatography of tryptic peptides, fractions A, B and C from Sephadex G-25*
The column size was 0.9 cm × 26 cm at 3°C, flow rate 12 ml/h, with stepwise increase in sodium acetate concentration at pH4.0. Vertical broken lines indicate the fractions that were combined.